

Acquisition du langage & phonétique/phonologie : outils et données

Yvan Rose
Memorial University of Newfoundland

Ce cours portera sur des aspects méthodologiques, technologiques et empiriques touchant la recherche et la formation académique dans le domaine de l'acquisition de la phonétique et de la phonologie des langues premières ainsi que des langues subséquentes.

Nous examinerons tout d'abord certaines traditions méthodologiques de la recherche sur l'acquisition de la phonétique et de la phonologie, notamment en établissant une distinction entre les approches expérimentales, naturelles et leurs respectives sous-catégories. Ce survol sera accompagné d'une description de certaines des méthodes les plus employées pour la capture de données d'acquisition. Bien que l'emphase sera sur la recherche portant sur des données de production, certaines méthodes utilisées pour la recherche en perception du langage seront aussi décrites. Cette exploration des méthodes d'investigation mettra au jour un aperçu de ce qu'elles permettent d'accomplir et, aussi, de certaines de leurs limites.

Cette discussion servira de préambule à une description des principaux outils informatiques couramment utilisés pour la recherche en acquisition de la phonologie. Ces outils incluent la version 4th Dimension de ChildPhon (Fikkert 1994, Levelt 1994, Freitas 1997), la version FileMaker Pro de ChildPhon (Rose 2003), Praat (www.fon.hum.uva.nl/praat/), Speech Filing System (<http://www.phon.ucl.ac.uk/resource/sfs/>), CLAN (<http://chilides.psy.cmu.edu/clan>), LIPP (<http://www.ihsys.com>) et Phon (<http://chilides.psy.cmu.edu/phon>). Un inventaire des principaux avantages et inconvénients de chacun de ces outils sera offert, avec une emphase sur les besoins les plus importants pour faciliter la recherche et la formation académique.

Le nouveau logiciel Phon, lequel a été lancé officiellement en novembre dernier (Rose et al. 2006), sera discuté plus en détail. Une esquisse des fonctions offertes par ce logiciel sera d'abord offerte, suivie d'un résumé des développements en cours d'implémentation.

Cette discussion sera suivie d'un exposé de l'inventaire des corpus de données existants et/ou disponibles publiquement, ainsi qu'une brève description de certains corpus toujours en construction. Nous commenterons sur les aspects forts et les lacunes de la base empirique actuelle. Notre discussion portera entre autres sur la nature des corpus existants, les méthodes utilisées pour leur élaboration, et le format dans lesquels ils ont été encodés. Comme nous le verrons, cet inventaire, actuellement limité tant de par sa taille que du point de vue de la diversité linguistique qu'il documente, est appelé à se développer de manière significative dans les prochaines années.

Cette portion de l'exposé sera fait en référence au nouveau projet PhonBank (MacWhinney et Rose, 2005) récemment mis en place pour favoriser l'entreposage et le partage des données phonologiques d'acquisition. Une invitation sera lancée à participer à cette initiative, laquelle permettra, à terme, de solutionner des problèmes méthodologiques et empiriques actuels et, ainsi, d'ouvrir de nouvelles avenues de recherche.